HER2 status predicts for upfront AI benefit: A TRANS-AIOG meta-analysis of 12,129 patients from ATAC, BIG 1-98 and TEAM with centrally determined HER2

Bartlett, John M.S.; Ahmed, Ikhlaaq; Regan, Meredith M.; Sestak, Ivana; Mallon, Elizabeth A.; Dell'Orto, Patrizia; Thürlimann, Beat; Seynaeve, Caroline; Putter, Hein; Van de Velde, Cornelis J.H.; Brookes, Cassandra L.; Forbes, John F.; Viale, Giuseppe; Cuzick, Jack; Dowsett, Mitchell; Rea, Daniel

DOI:
10.1016/j.ejca.2017.03.033

Original Research Article:

HER2 status predicts for upfront AI benefit: a TRANS-AIOG meta-analysis of 12129 patients from ATAC, BIG 1-98 and TEAM with centrally determined HER2.

John M. S. Bartlett[a,b], Ikhlaaq Ahmed[c], Meredith M. Regan[d], Ivana Sestak[e], Elizabeth A. Mallon[f], Patrizia Dell'Orto[g], Beat Thürlimann[h], Caroline Seynaeve[i], Hein Putter[j], Cornelis J. H. Van de Velde[k], Cassandra L. Brookes[l], John F. Forbes[m], Giuseppe Viale[g], Jack Cuzick[e], Mitchell Dowsett[n], Daniel W. Rea[c] on behalf of the Translational Aromatase Inhibitor Overview Group (Trans-AIOG).

Affiliations:

[a]Transformative Pathology, Ontario Institute for Cancer Research (OICR), MaRS Centre, 661 University Avenue, Suite 510, Toronto, ON M5G 0A3, Canada; john.bartlett@oicr.on.ca; [b]University of Edinburgh Cancer Research Centre, Institute of Genetics & Molecular Medicine, The University of Edinburgh, Western General Hospital, Crewe Road South, Edinburgh, EH4 2XR, United Kingdom (UK); john.bartlett@oicr.on.ca; [c]Cancer Research UK Clinical Trials Unit, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK; I.Ahmed.2@bham.ac.uk; d.w.rea@bham.ac.uk; [d]International Breast Cancer Study Group Statistical Center, Dept. of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Harvard Medical School, 450 Brookline Avenue, Boston, MA 02215, United States (US); mregan@jimmy.harvard.edu; [e]Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University, Charterhouse Square, London, EC1M 6BQ, UK; i.sestak@qmul.ac.uk; j.cuzick@qmul.ac.uk; [f]Department of Pathology, Elizabeth University Hospital, 1345 Govan Rd, Govan, Glasgow, G51 4TF, UK; Elizabeth.Mallon2@ggc.scot.nhs.uk; [g]Department of Pathology, European Institute of Oncology and University of Milan, Via Giuseppe Ripamonti, 435, 20141 Milan, Italy; patrizia.dellorto@ieo.it; giuseppe.viale@ieo.it; [h]Breast Center, Kantonsspital, St. Gallen, and Swiss Group for Clinical Cancer Research SAKK, CH-9007 St. Gallen, Switzerland; beat.thuerlimann@kssg.ch; [i]Department of Medical Oncology, Erasmus MC Cancer Institute, PO Box 5201, Rotterdam, 3008 AE, The Netherlands; c.seynaeve@erasmusmc.nl; [j]Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands; h.putter@lumc.nl; [k]Division of Surgery, Leiden University Medical Center, PO Box 9600, 2300 RC, Leiden, The Netherlands; c.j.h.van_de_velde@lumc.nl; [l]Leicester Clinical Trials Unit, University of Leicester, Leicester, LE5 4PW, UK; cassey.brookes@leicester.ac.uk; [m]Surgical Oncology, University of Newcastle, Calvary Mater Newcastle Hospital, Locked Bag 7, HRMC (Newcastle) NSW 2310, Australia; john.forbes@anzbctg.org; [n]Academic Department of Biochemistry, Royal Marsden Hospital, and Endocrinology, Breakthrough Breast Cancer Centre, Institute of Cancer Research, Fulham Road, London, SW3 6JJ, UK; Mitchell.Dowsett@icr.ac.uk.

Corresponding Author:

Prof. John M. S. Bartlett
Transformative Pathology
Ontario Institute for Cancer Research
MaRS Centre
661 University Avenue
Suite 510
Toronto, ON M5G 0A3
Canada
1-647-259-4251
John.bartlett@oicr.on.ca

Funding

The BIG 1-98 trial was financed by Novartis and coordinated by the IBCSG [clinicaltrials.gov IDs=NCT00004205], including the design of the trial, data management, medical review, pathology review, and statistical support. The ATAC Trial was funded by Astra Zeneca [clinicaltrials.gov IDs=NCT00849030] and Cancer Research UK. The TEAM trial was funded by Pharmacia/Pfizer [clinicaltrials.gov IDs=NCT00279448] and aspects of the biomarker work funded by Cancer Research UK. John Bartlett was supported by funding from OICR which is provided through the Ontario Ministry of Research, Innovation and Science.

Abstract

***Background*:** A meta-analysis of the effects of HER2 status, specifically within the first 2-3 years of adjuvant endocrine therapy, has the potential to inform patient selection for upfront AI therapy or switching strategy tamoxifen followed by AI. The pre-existing standardization of methodology for HER2 (IHC/FISH) facilitates analysis of existing data for this key marker.

Methods: Following a prospectively designed statistical analysis plan, patient data from 3 phase III trials (ATAC, BIG-1-98 and TEAM) comparing an AI to tamoxifen during the first 2-3 years of adjuvant endocrine treatment were collected and a treatment-by-marker analysis of distant recurrence-free interval–censored at 2-3 years treatment–for HER2 status x AIvsTam treatment was performed to address the clinical question relating to efficacy of "upfront" versus "switch" strategies for AIs.

Results: A prospectively planned, patient-level data meta-analysis across 3 trials demonstrated a significant treatment (AIvsTam) by marker (HER2) interaction in a multivariate analysis; (interaction HR=1.61, 95%CI 1.01-2.57; $p<0.05$). Heterogeneity between trials did not reach statistical significance. The HER2-ve group gained greater benefit from AI versus Tam (HR=0.70, 95%CI 0.56-0.87) than the HER2+ve group (HR=1.13, 95%CI 0.75-1.71). However, the small number of HER2+ve cases (n=1092 across the 3 trials) and distant recurrences (n=111) may explain heterogeneity between trials.

Conclusions: A patient level data meta-analysis demonstrated a significant interaction between HER2 status and treatment with AIvsTam in the first 2-3 years of adjuvant endocrine therapy. Patients with HER2-ve cancers experienced improved outcomes (distant relapse) when treated with upfront AI rather than tamoxifen, whilst patients with HER2+ve cancers fared no better or slightly worse in the first 2-3 years. However, the small number of HER2+ve cancers/events may explain a large degree of heterogeneity in the HER2+ve groups across all 3 trials. Other causes, perhaps related to subtle differences between AIs, cannot be excluded and warrant further exploration.

Keywords

Breast cancer, HER2, prediction, aromatase inhibitor, meta-analysis

Introduction

For well over 20 years the HER2 (neu/c-erb-b2) oncogene has been associated with resistance to endocrine therapy [1]. As knowledge relating to extended type I receptor tyrosine kinase family (RTK; EGFr, HER2, HER3 and HER4) signaling was developed, functional and clinical evidence substantiating the link between resistance to tamoxifen therapy and type I RTK expression became more extensive [1-3]. A decade ago we suggested that analysis of type I RTKs might be of value in determining which patients were most likely to benefit from aromatase inhibitor (AI) rather than tamoxifen therapy [4]. At this time we made two critical observations relevant for the clinical setting, firstly, that the impact of HER2 and other type I RTK status on outcome following tamoxifen therapy was time dependent, and secondly that HER2 was not the sole driver of tamoxifen resistance in early breast cancer [4, 5].

The type I RTK family (HER1-4) form 10 homo- or heterodimers and are activated by a broad range of ligands leading to a complex inter-relationship between signaling kinases and downstream pathways [3]. There is evidence that HER4 does not promote breast cancer proliferation in vivo and is linked to good prognosis in breast cancer patients [4, 6]. In contrast breast tumors expressing HER1, HER2 or HER3 receptors exhibit increased proliferation in vivo and are associated with poor outcome [4].

Establishing the impact of specific genes on cancer prognosis in the clinical setting is complicated by multiple factors including; the impact of multimodal therapy (surgery, radiotherapy, chemotherapy, endocrine therapy), the size and consistency of the patient cohort and the frequency of marker expression. These challenges are recognized and addressed by existing guidelines (REMARK) [7]. Several studies have demonstrated time dependency of both molecular and clinical prognostic features of breast cancer [8, 9]. Both the time dependency of the impact of type I RTKs and the inter-relationship between different receptors in the RTK family require consideration in any approach to the determination of the prognostic or predictive impact of these markers in early breast cancer.

At the time of our 2005 study [4], results from the first clinical trials of AIs demonstrated the superiority of AIs regarding disease-free survival (DFS) when compared with tamoxifen for postmenopausal, estrogen receptor (ER) positive early breast cancer [10]. When linked to preclinical evidence that type I RTKs mediate resistance to tamoxifen, in part through intracellular signaling [11] and post-translational modification of ER [1, 12], this led to a hypothesis that clinical resistance to tamoxifen treatment mediated via HER1-3 expression and signaling could be circumvented/reversed by the clinical use of AIs [13]. Evidence from three trials (ATAC, BIG-1-98 and TEAM) [14-16] in which central HER2 testing was performed [17-19] provided support for an interaction between HER2 expression and response to AIs versus tamoxifen. However, the effect observed suggested that, contrary to earlier hypotheses, HER2 negative (HER2-ve) cancers derived maximal benefit from AIs whilst HER2 positive (HER2+ve) cancers appeared to exhibit a generalized endocrine resistance and might, in some trials, even perform worse on AIs rather than tamoxifen. However, none of the translational studies within these trials were, individually, statistically powered to provide high level evidence for the suggested interaction between endocrine therapy and HER2 which might impact clinical practice.

This challenge led to the initiation, through the Translational sub-group of the Aromatase Inhibitor Overview Group (Trans-AIOG) of the current meta-analysis of data from 3 randomised trials (ATAC, BIG-1-98 and TEAM). No HER1/3 data was available from two studies limiting the current analysis to testing the interaction between HER2 and AI versus tamoxifen treatment in the first 2-3 years of treatment using a two-sided statistical test. This interval was chosen as most likely to impact clinical decisions relating to the initiation of early AI therapy versus the use of a switching strategy (Tam→AI). In addition, this analysis plan may account for time dependent effects of HER2 observed in multiple studies and allows inclusion of patients from both switching and continuous treatment arms maximizing statistical power. We report the results of this prospectively planned analysis designed according to the guidelines produced by Simon et al [20], to produce level IB evidence relating to the proposed treatment by marker effect.

Materials and Methods

A predefined statistical analysis plan to determine by individual patient data (IPD) meta-analysis the use of HER2 as a biomarker for selection of upfront AIs compared to tamoxifen in the first 2-3 years of treatment in postmenopausal early breast cancer patients was developed and approved by the Trans-AIOG core investigators (JMSB, DR, GV, MMR, IS, CLB, IA; see supplementary data). The hypothesis to be tested was that HER2 is a predictive biomarker for AI benefit in HER2–ve patients during the first 2-3 years of endocrine therapy, i.e. that benefit from AIs over tamoxifen in that time period is not seen in patients with HER2+ve disease but is confined to patients with HER2-ve disease.

The primary outcome was defined as distant recurrence-free interval (DRFI) up until the (pre-planned) 2-3 year treatment switch time (for trials with a switching element). DRFI was measured in days for all patients from randomisation date until earliest documentation of distant recurrence, or was censored at the visit date when treatment switch occurred, or would have occurred if assigned to do so (around 2-3 years depending on trial protocols), or at last follow-up or death without distant recurrence, whichever was earliest. A secondary outcome analysis was invasive disease-free survival (IDFS) up until the 2-3 year treatment switch time to be measured in days for all patients from randomisation date until earliest documentation of recurrent disease to include any invasive ipsilateral recurrence, invasive contralateral disease, loco-regional or distant recurrence, any second primary (non-breast) malignancy, or death from any cause. Patients alive and in follow-up without evidence of disease were censored at the visit date when treatment switch (around 2-3 years) occurred, or would have occurred or last follow-up, whichever was earlier.

Data collection and audit

Following approval of the SAP by each translation science group the IPD analysis used data from patients randomised from three trials TEAM [14], ATAC [15] and BIG 1-98 [16]. All biomarker data included in this meta-analysis have been previously published [17-19]. Information was sought for every eligible patient included in the sub-study including randomisation date, age at enrollment, histological grade, stage, nodal

information, tumour size, adjuvant and neoadjuvant treatment, disease recurrence (including dates of first local, regional, distant, new primary disease), follow-up status and time, survival status and cause of death.

Sample Size

Using sample size calculations methods proposed by Schmoor [21], with a two-sided alpha of 0.05 and assuming an interaction hazard ratio (HR) of 2.44, 5.8% event rate, and HER2-positive prevalence of 10%, a sample size of 12448 patients would give greater than 90% power to detect a treatment–biomarker interaction within the 2-3 year treatment period. The sample size was calculated using the 'powerEpiInt2' function in the 'powerSurvEpi' [22] package in R (R Foundation for Statistical Computing, Vienna, Austria).

HER2/ER/PgR positivity

HER2 status was determined centrally and recorded as either positive or negative in each study according to published protocols [17-19] and complied with the ASCO-CAP guidelines updated in 2013 [23].

All three participating trials assessed estrogen receptor (ER) and progesterone receptor (PgR) status centrally using different staining methodologies for each trial [17, 18, 24, 25]. For all eligible patients, ER positivity was defined as ≥1%, similarly PgR-positivity was defined ≥1% according to ASCO-CAP guidelines [26].

Statistical methods

Individual patient-level data (IPD) were collected centrally in Birmingham, UK. The statistical analysis was performed on an intention-to-treat basis within each of the three trials. Patients were excluded from the analysis if tumors were not ER-positive on central testing (<1% immunoreactive cells or ER not determined) or if HER2 status was not centrally determined. An IPD meta-analysis of HER2 as a predictive biomarker for response to AI versus tamoxifen treatment was conducted. Log hazard ratios (HRs) were calculated (separately for HER2-ve and HER2+ve) for each trial taking into account other important prognostic factors (age at enrollment, grade, size, nodal status) and then combined using random-effect inverse variance meta-analysis to give an overall treatment-by-biomarker effect estimate [27]. The interaction term between HER2 and the treatment assignment was calculated in each trial separately, and then pooled together across the three trials using a random-effects meta-analysis. This analysis was performed for both the primary and secondary endpoints.

Results

In total 23,669 women were randomized to the TEAM (9779), ATAC (5880) and BIG-1-98 (8010) clinical trials (Supplementary Table 1). Collectively 13,336 tissue samples were received at the core pathology laboratories, and 12,671 successfully stained for the biomarkers ER, PgR and HER2. Translational data from all 3 trials were collected centrally and following central checking 542 cases were excluded from the analysis because they were ER negative or unknown or had missing HER2 status. In total 12,129 cases (97.4% of the planned sample size) were available for this analysis (Supplementary Table 1). Overall 1092 (9%) patients had HER2 positive disease, ranging from 6.3% to 12.3% across the three trials (Supplementary Table 2).

Differences between the trial cohorts with respect to clinicopathological variables reflect the recruitment strategies for each trial. Within each trial the translational cohorts have been previously shown to represent the main populations [17, 18, 24, 25] (Supplementary Table 2). Overall, 473/12129 patients (3.9%) experienced a distant recurrence within the 2-3 year period defined for the analysis, and 892/12129 (7.4%) patients experienced an IDFS event (Supplementary Table 3).

Treatment-by-HER2 interaction

The prospectively-planned IPD meta-analysis across 3 phase III trials demonstrated a significant treatment (AI vs. tamoxifen) by marker (HER2 status) interaction in both univariate (interaction HR=1.76, 95%CI 1.14-2.71; $p<0.05$) and multivariate analysis (interaction HR=1.61, 95%CI 1.01-2.57; $p<0.05$). These pooled results for the unadjusted and adjusted HER2 and treatment interactions are similar, with both HRs greater than 1 (Figure 1A,1B) confirming a statistically-significant difference in the treatment effect (AI versus tamoxifen) between HER2 positive and HER2 negative cancers. Whilst there is evidence of heterogeneity for this effect between trials (I squared of approximately 60%) this effect does not reach statistical significance.

Overall, in line with the treatment-by-marker interaction result, the HER2-ve group gained greater benefit from AI versus tamoxifen in both the unadjusted analyses (Figure 2A) and adjusted analyses for clinicopathological variables (adjusted HR=0.70, 95%CI 0.56-0.87) than the HER2+ve group (HR=1.13, 95%CI 0.75-1.71) (Figure

2B). For HER2-ve cases similar HRs were observed in all three trials with no significant heterogeneity (adjusted analyses, I squared=0%, p=0.587). However, significant heterogeneity was observed between trials for the HER2+ve subgroup (I-squared=70.8% p<0.05; Figure 2B). The small number of HER2+ve cases (1092/12129 across 3 trials) and distant recurrences (111 for HER2+ve cancers across all trials) may in part explain the heterogeneity between trials for this effect. Overall for the 12,129 patients included in this study with centrally confirmed ER and HER2 status, treatment with AIs resulted in a 22% reduction in risk of distant recurrence when compared to tamoxifen (HR=0.78, 95%CI 0.64-0.94) when results were pooled across all cases (HER2 +ve and HER2-ve).

Trial by trial subgroups

Individual analyses within each trial cohort confirmed a significant treatment-by-HER2 effect within the TEAM trial population, in both unadjusted (HR=2.66; 95%CI 1.47-4.82, p<0.001; Supplementary Table 4) and adjusted (HR=2.75, 95%CI 1.38-5.48, p=0.004; Supplementary Table 5) analyses. However, within the ATAC trial the treatment-by-marker effect did not reach statistical significance in either the unadjusted (HR=1.87; 95%CI 0.62-5.65; p>0.05; Supplementary Table 6) or adjusted analyses (HR=1.44, 95%CI 0.46-4.96; p>0.05; Supplementary Table 7). Finally, the BIG-1-98 trial showed no evidence for a treatment-by-marker interaction in either unadjusted (HR=0.86, 95%CI 0.40-1.85; Supplementary Table 8) or adjusted analyses (HR=0.87, 95%CI 0.40-1.87; Supplementary Table 9).

Invasive Disease-free Survival

A secondary pre-planned analysis of IDFS showed similar results in both adjusted and unadjusted analyses with a statistically significant (p<0.05) treatment-by-HER2 interaction (Figure 3A,3B). However, there was statistically significant heterogeneity in the treatment-by-marker effect between different trials (p<0.05, Figure 3A,3B). Results are again similar for the unadjusted and adjusted treatment effects in the two HER2 subgroups (Figure 3). The treatment effects for the HER2-ve group are statistically significant in favour of AI treatment (HR=0.77, 95%CI. 0.67-0.89) with minimal heterogeneity between the 3 trial cohorts (Figure 3B). Whilst the

pooled treatment effect for the HER2+ve group is greater than 1 (favouring tamoxifen) the 95%CI crosses 1 (Figure 3A, 3B) and there is significant heterogeneity between trials (I squared>80%, p<0.05).

Discussion

In this preplanned individual patient-level data meta-analysis across 3 pivotal randomized trials (ATAC, BIG-1-98 and TEAM) using data from 12,129 postmenopausal, ER+ve early breast cancer patients with centrally determined ER and HER2 status, we demonstrated a significant interaction between HER2 status and treatment with AIs versus tamoxifen in the 2-3 years of adjuvant endocrine therapy (prior to the clinically-relevant point of "switching" between tamoxifen and AIs) (interaction HR=1.61, 95%CI 1.01-2.57, p<0.05). Patients with HER2-ve cancers experienced improved outcomes (30% reduction in distant relapse risk) when treated with an AI whilst patients with HER+ve cancers fared no better, or slightly worse during AI treatment in this period. This prospectively planned and statistically powered analysis achieved 97.4% of the prospectively planned sample size, was compliant with both BRISQ and REMARK guidelines [7, 28] and satisfies the required criteria for level 1B evidence according to Simon et al [20]. Whilst heterogeneity was observed in the HER2 positive subgroup, which represents under 10% of the entire population, this heterogeneity did not reach statistical significance within the primary endpoint population. In all aspects this analysis satisfies the current standards for a practice changing validation of a predictive biomarker in early breast cancer [20].

However, as with all such studies, results must be applied in the context of existing breast cancer treatment options available to patients. Critically, for the majority of HER2+ve cases Herceptin/trastuzumab™ treatment was not available in the absence of chemotherapy at the time of conduct of the 3 trials used in this analysis. Therefore, for HER2 positive breast cancers the potential impact of HER2-directed therapies remains unknown, this alone precludes recommendations to change practice. Furthermore, whilst this study might be interpreted to suggest that patients with HER2+ve cancers derive no benefit in terms of freedom from distant recurrence from upfront AI (versus tamoxifen) treatment (within the first 2-3 years), we recognize that this study includes heterogeneous results from different studies which preclude recommendations to change treatment. Importantly,

the small number of HER2+ve cancers and recurrences in even this large analysis may explain the large degree of heterogeneity in the HER2+ve subgroups across all 3 trials. Rates of HER2 positivity varied between trials (6.3%, 10.6% and 12.3% in BIG-1-98, ATAC and TEAM respectively) despite all 3 trials using ASCO-CAP guidelines for HER2 testing [29]. Almost 50% of the HER2+ve cancers and over 50% of distant recurrences in the HER2 positive subgroup were from a single study, the TEAM trial, which was the only study to use exemestane, an irreversible, non-competitive AI. Furthermore, these results only apply to treatment and events occurring during the first 2-3 years of treatment (pre-switching) and do not provide information on longer term patient outcome. These results contrast with the recent meta-analysis from the EBCTCG where no interaction between HER2 and treatment was observed. The EBCTCG analyses, however, combined trials with switching strategies and those comparing 5 years of AI versus tamoxifen [30], and used locally- rather than centrally-determined ER and HER2 results testing. Further, within the overview over 70% of cases had "unknown" HER2 status reflecting the lack of data in this field. In addition, emerging evidence of activity for both EGFr/HER2 as candidate predictive biomarkers, although currently restricted to a single study [17], suggest additional research is required using a broader approach to determining the effects of type I RTK signaling.

This meta-analysis provides further evidence for differential benefit from upfront AI versus tamoxifen treatment between HER2+ve and HER2-ve ER positive postmenopausal early breast cancer patients. This study supports the positive impact of upfront AI in HER2-ve cancers. It raises the challenging possibility that treatment of HER2+ve luminal breast cancers with AI may, at best, provide no additional benefit over tamoxifen, and at worst be detrimental. However, we cannot at this time provide irrefutable evidence for such a detrimental effect especially in the context of HER2-targeted therapies. Therefore, on the basis of this study it is not appropriate to suggest excluding HER2 positive breast cancers from early treatment with aromatase inhibitors. Nor do we propose changes to current adjuvant endocrine therapy guidelines on the basis of HER2 status. The observed heterogeneity between trials, coupled with minimal treatment using HER2 directed therapies in this group precludes recommendations to change patient management at this time. However, the observed effect, coupled with previous data from the TEAM study [17] (combining EGFr/HER2/HER3) provides strong support for

further research in this field, aiming to further knowledge on optimal endocrine therapy for subgroups of ER+ve postmenopausal breast cancer patients.

Acknowledgements

The ATAC, TEAM and International Breast Cancer Study Group (IBCSG) collaborators are particularly indebted to the patients, pathologists, physicians, nurses, and data managers who participated in the respective trials, and staff at the centers and cooperative groups of the ATAC, TEAM and BIG 1-98 Collaborative Groups. The BIG 1-98 trial was financed by Novartis and coordinated by the IBCSG [clinicaltrials.gov IDs=NCT00004205], including the design of the trial, data management, medical review, pathology review, and statistical support. The ATAC Trial was funded by Astra Zeneca [clinicaltrials.gov IDs=NCT00849030] and Cancer Research UK. The TEAM trial was funded by Pharmacia/Pfizer [clinicaltrials.gov IDs=NCT00279448] and aspects of the biomarker work funded by Cancer Research UK. John Bartlett was supported by funding from OICR which is provided through the Ontario Ministry of Research, Innovation and Science. The funding sources of the study had no role in the design of the study, the collection, analysis, or interpretation of the data, the writing of this report, or in the decision to submit the article for publication.

Conflict of Interest statement

Beat Thürlimann reports stock ownership of Novartis. Jack Cuzick reports that AstraZeneca supported the ATAC study. All remaining authors have declared no conflicts of interest.

Figure Legends

Figure 1:

Forest plots for DFRI biomarker and treatment interaction.

Panel A = unadjusted/univariate. Panel B = adjusted/multivariate analysis.

Figure 2:

Treatment (AI versus Tam) effect by marker (HER2) by trial and combined.

Panel A = unadjusted/univariate analysis. Panel B = adjusted/multivariate analysis.

Figure 3:

Treatment (AI versus Tam) effect by marker (HER2) by trial and combined for DFS (secondary endpoint).

Panel A = unadjusted/univariate analysis. Panel B = adjusted/multivariate analysis.

References

[1] Benz CC, Scott GK, Sarup JC, et al. Estrogen-Dependent, Tamoxifen-Resistant Tumorigenic Growth of Mcf-7 Cells Transfected with Her2/Neu. *Breast Cancer Res Treat* 1992,**24**(2), 85-95.

[2] Ross JS, Fletcher JA. The HER-2/neu oncogene in breast cancer: prognostic factor, predictive factor, and target for therapy. *Stem Cells* 1998,**16**(6), 413-418.

[3] Yarden Y, Pines G. The ERBB network: at last, cancer therapy meets systems biology. *Nat Rev Cancer* 2012 Aug,**12**(8), 553-563.

[4] Tovey SM, Dunne B, Witton CJ, Forsyth A, Cooke TG, Bartlett JMS. Can Molecular Markers Predict When to Implement Treatment with Aromatase Inhibitors in Invasive Breast Cancer? *Clin Cancer Res* 2005 Jul 1,**11**(13), 4835-4842.

[5] Witton CJ, Reeves JR, Going JJ, Cooke TG, Bartlett JMS. Expression of the HERI-4 family of receptor tyrosine kinases in breast cancer. *Journal of Pathology* 2003,**200**(3), 290-297.

[6] Naresh A, Long WW, Vidal GA, et al. The ERBB4/HER4 intracellular domain 4ICD is a BH3-only protein promoting apoptosis of breast cancer cells. *Cancer Res* 2006,**66**(12), 6412-6420.

[7] McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. REporting recommendations for tumour MARKer prognostic studies (REMARK). *British Journal of Cancer* 2005,**93**(4), 387-391.

[8] Ali AM, Provenzano E, Bartlett JM, et al. Prognosis of early breast cancer by immunohistochemistry defined intrinsic sub-types in patients treated with adjuvant chemotherapy in the NEAT/BR9601 trial. *Int J Cancer* 2013 Sep 15,**133**(6), 1470-1478.

[9] Stephen J, Murray G, Cameron DA, et al. Time dependence of biomarkers: non-proportional effects of immunohistochemical panels predicting relapse risk in early breast cancer. *Br J Cancer* 2014 Dec 9,**111**(12), 2242-2247.

[10] Dowsett M, Cuzick J, Ingle J, et al. Meta-Analysis of Breast Cancer Outcomes in Adjuvant Trials of Aromatase Inhibitors Versus Tamoxifen. *Journal of Clinical Oncology* 2010,**28**(3), 509-518.

[11] McGlynn LM, Tovey S, Bartlett JM, Doughty J, Cooke TG, Edwards J. Interactions between MAP kinase and oestrogen receptor in human breast cancer. *Eur J Cancer* 2013 Apr,**49**(6), 1176-1186.

[12] Gee JM, Robertson JF, Gutteridge E, et al. Epidermal growth factor receptor/HER2/insulin-like growth factor receptor signalling and oestrogen receptor activity in clinical breast cancer. *Endocr Relat Cancer* 2005,**12**, S99-S111.

[13] Ellis MJ. Importance of correlative science in advancing hormonal therapy and a new clinical paradigm for neoadjuvant therapy. *Annals of Surgical Oncology* 2004,**11**(1), 9S-17S.

[14] van de Velde CJH, Rea D, Seynaeve C, et al. Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial. *Lancet* 2011,**377**(9762), 321-331.

[15] Howell A, Cuzick J, Baum M, et al. Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. *Lancet* 2005,**365**(9453), 60-62.

[16] Thurlimann B, Keshaviah A, Coates AS, et al. A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer. *N Engl J Med* 2005 Dec 29,**353**(26), 2747-2757.

[17] Bartlett JM, Brookes CL, Piper T, et al. Do type 1 receptor tyrosine kinases inform treatment choice? A prospectively planned analysis of the TEAM trial. *Br J Cancer* 2013 Oct 29,**109**(9), 2453-2461.

[18] Cuzick J, Dowsett M, Wale C, et al. Prognostic Value of a Combined ER, PgR, Ki67, HER2 Immunohistochemical (IHC4) Score and Comparison with the GHI Recurrence Score - Results from TransATAC. *Cancer Res* 2009,**69**(24), 503S.

[19] Rasmussen BB, Regan MM, Lykkesfeldt AE, et al. Adjuvant letrozole versus tamoxifen according to centrally-assessed ERBB2 status for postmenopausal women with endocrine-responsive early breast cancer: supplementary results from the BIG 1-98 randomised trial. *The Lancet Oncology* 2008,**9**(1), 23-28.

[20] Simon R. Development and validation of biomarker classifiers for treatment selection. *Journal of Statistical Planning and Inference* 2008,**138**(2), 308-320.

[21] Schmoor C, Sauerbrei W, Schumacher M. Sample size considerations for the evaluation of prognostic factors in survival analysis. *Statistics in Medicine* 2000,**19**, 441-452.

[22] Qui W, Chavarro J, Lazarus R, Rosner B, Ma J. Power and sample size calculation for survival analysis of epidemiological studies. R package version 0.0.6. 2012.

[23] Wolff AC, Hammond ME, Hicks DG, et al. Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Update. *Journal of Clinical Oncology* 2013 Oct 7,**31**(31), 3997-4013.

[24] Bartlett JMS, Brookes CL, Robson T, et al. Estrogen Receptor and Progesterone Receptor As Predictive Biomarkers of Response to Endocrine Therapy: A Prospectively Powered Pathology Study in the Tamoxifen and Exemestane Adjuvant Multinational Trial. *Journal of Clinical Oncology* 2011,**29**(12), 1531-1538.

[25] Viale G, Regan MM, Maiorano E, et al. Prognostic and Predictive Value of Centrally Reviewed Expression of Estrogen and Progesterone Receptors in a Randomized Trial Comparing Letrozole and Tamoxifen Adjuvant Therapy for Postmenopausal Early Breast Cancer: BIG 1-98. *Journal of Clinical Oncology* 2007 Sep 1,**25**(25), 3846-3852.

[26] Hammond MEH, Hayes DF, Dowsett M, et al. American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations for Immunohistochemical Testing of Estrogen and Progesterone Receptors in Breast Cancer (Unabridged Version). *Archives of Pathology & Laboratory Medicine* 2010,**134**(7), E48-E72.

[27] Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *Bmj* 2011,**342**, d549.

[28] Moore HM, Kelly AB, Jewell SD, et al. Biospecimen reporting for improved study quality (BRISQ). *J Proteome Res* 2011 Aug 5,**10**(8), 3429-3438.

[29] Wolff AC, Hammond ME, Hicks DG, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol* 2013 Nov 1,**31**(31), 3997-4013.

[30] Dowsett M, Forbes JF, Bradley R, et al. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. *Lancet* 2015 Oct 3,**386**(10001), 1341-1352.